Epidemic variability in complex networks

Pascal Crépey,[1,2] Fabián P. Alvarez,[1,2] and Marc Barthélemy[3,4]

[1]INSERM, Unité de Recherche "Épidémiologie, Systèmes d'Information et Modélisation" (U707), Paris, F-75012, France
[2]Université Pierre et Marie Curie, Faculté de Médecine Pierre et Marie Curie,
Unité de Recherche "Épidémiologie, Systèmes d'Information et Modélisation", Paris, F-75012, France
[3]CEA-Centre d'Etudes de Bruyères-le-Châtel, Département de Physique Théorique et Appliquée
BP12, 91680 Bruyères-Le-Châtel, France
[4]School of Informatics, Indiana University, Eigenmann Hall,
Bloomington, IN 47408, USA
(Dated: 6th February 2008)

We study numerically the variability of the outbreak of diseases on complex networks. We use a SI model to simulate the disease spreading at short times in homogeneous and in scale-free networks. In both cases, we study the effect of initial conditions on the epidemic dynamics and its variability. The results display a time regime during which the prevalence exhibits a large sensitivity to noise. We also investigate the dependence of the infection time of a node on its degree and its distance to the seed. In particular, we show that the infection time of hubs have non-negligible fluctuations which limit their reliability as early-detection stations. Finally, we discuss the effect of the multiplicity of paths between two nodes on the infection time. In particular, we demonstrate that the existence of even long paths reduces the average infection time. These different results could be of use for the design of time-dependent containment strategies.

PACS numbers: 89.75.Hc, 87.23.Ge, 87.19.Xx

I. INTRODUCTION

Many complex systems display a very heterogeneous degree distribution [1, 2, 3, 4] characterized by a power law decay of the form $P(k) \sim k^{-\gamma}$. This form implies the absence of a characteristic scale hence the name of "scale-free network" (SFN) [5, 6]. Among these networks, a certain number are of a great interest to epidemiology [4, 7, 8] and it is thus very important to understand the effect of their topology on the spreading dynamics of a disease. One of the most relevant results is that disease spreading does not show an endemic threshold in SFN when the population size is infinite and $\gamma \leq 3$ [9, 10, 11, 12, 13]. This result means that a disease propagates very easily on a large SFN whatever the value of its transmission probability. In addition, recent studies showed that the presence of hubs in SFN not only facilitates the spread of a disease but also accelerates dramatically its outbreak [14, 15, 17].

The long-tailed degree distribution of SFN is the signature of the presence of a non-negligible number of highly connected nodes. These hubs were already identified in the epidemiological literature as superspreaders [18, 19]. Consequently, from a public health point of view, studying the spreading of epidemics on SFN is all the more appropriate. Superspreading events affect the basic reproductive number R_0—a widely used epidemiological parameter [19, 21]—making its estimate from real-world data difficult [22, 23, 24]. As a matter of fact, it seems that superspreading events appeared in the onset of the recent SARS outbreak [23, 24, 25, 26] and could be crucial for the new emergent diseases and bioterrorist threats. Their potential threat justifies detailed studies of the incidence of the degree distribution at the initial stage of epidemics.

The variability plays an important role in the accuracy and the forecasting capabilities of numerical models and has thus to be quantified in order to assess the meaningfulness of simulations with respect to real outbreaks [20]. Using a numerical approach, we analyze the evolution of epidemics generated by different sets of initial parameters, both for SFN and homogeneous random networks (RN). We use the Barabási-Albert model (BA) [1] for generating a SFN and the Erdös-Renyi network (ER) [27] as a prototype for RN. Concerning the epidemic modeling, a simple and classical approach is to consider that individuals are only in two distinct states, infected (I) or susceptible (S). There is initially a number of $i_0 N$ infected individuals and any infected node can pass the disease to his neighbors [19, 21]. The probability per unit time to transmit the disease—the spreading rate—is denoted by λ and once a susceptible node is infected it remains in this state. In more elaborated models, an infected individual can change its state to another category, for example, coming back to susceptible (SIS), or going to immunized or dead (SIR) [19, 21]. This S $\to$ I approach (SI), in spite of its simplicity, is a good approximation at short times to more refined models such as the SIS or SIR models. The SI model on both SFN and RN is thus well adapted to the characterization of the variability of the initial stages of epidemic outbreaks spreading in complex networks, which is the focus of this article.

The outline of the paper is the following. In section II, we study the fluctuations of the prevalence and we identify different parameters controlling them. In particular, we highlight the effects due to different realizations of the network as well as different initial conditions. We also investigate the influence of the nodes degree on the prevalence variability. In section III, we present results on the infection time and its variation with the degree and with the distance from the origin of infections. We also discuss the effect of the number

of paths between two nodes on the infection time. Finally, we discuss our results and conclude in section IV.

II. PREVALENCE FLUCTUATIONS

A. Intra and inter-networks fluctuations

We analyze in this section the effect of the underlying network topology on the variability of outbreaks. It is indeed important to understand whether the local fluctuations of the structure of the network can have a large impact on the development of epidemics.

In order to analyze this effect, we measure the variability of outbreaks as the relative variation of the prevalence (density of infected individuals $i(t)$) given by

$$CV[i(t)] = \frac{\sqrt{\langle i(t)^2 \rangle - \langle i(t) \rangle^2}}{\langle i(t) \rangle}. \tag{1}$$

In order to evaluate this quantity we run simulations for different "model sets": first, for a given number of outbreaks on a single network, second for a single outbreak on different networks, and finally several outbreaks on different networks. We show in Fig. 1 the curves $CV[i(t)]$ computed for both the RN (thin lines) and the SFN (bold lines) and for two of these model sets: 10^3 outbreaks spreading on the same network (dashed curves), and a single outbreak per network for 10^3 different networks (plain curves). The curves representing these two model sets are nearly superimposed for both network topologies. The curves obtained from model sets made of 10 outbreaks on 100 networks and 100 outbreaks on 10 networks coincide with the other model sets (not shown in the figure). These results indicate that the contribution to the variability of $i(t)$ given by a particular network realization is essentially the same as the one generated by different outbreaks on the same network. This confirms the intuitive idea that sampling different parts of a large network is equivalent to average over different networks. Consequently, studying variability of epidemics simulated on one large enough network (intra-network) will lead practically to the same conclusions as studying variability on several instances of that network (inter-network). Furthermore, it means that the results described in the next sections for one network can be generalized to any instances of BA and ER networks.

Fig. 1 also reveals interesting facts about the time behavior of $CV[i(t)]$ on complex networks. Since the initial prevalence is fixed and is the same for all instances, CV is initially equal to zero and can only increase. At very large times, almost all nodes are infected implying that $\lim_{t \to \infty} CV = 0$. This argument implies the existence of a peak which—as shown in Fig. 1—is located for BA networks at the beginning of the outbreak, with a maximum value larger than the one obtained for ER networks. In order to characterize the relation between the variability peak and the network heterogeneity, we define τ_v as the time at which the maximum of $CV[i(t)]$ is reached. We also use the fact that the heterogeneity of the network degree—often quantified by $\kappa = \langle k^2 \rangle / \langle k \rangle$—is related to the
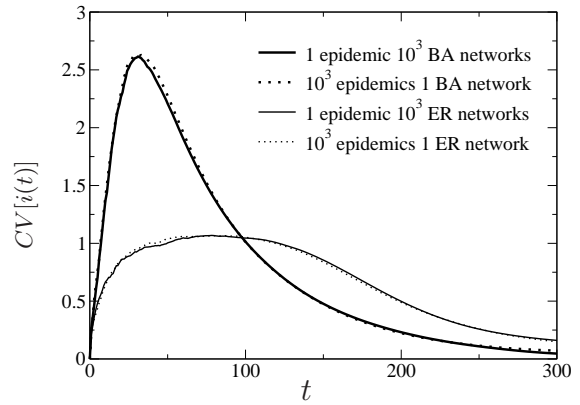

Figure 1: Evolution in time of the coefficient of variation of the density of infected ($CV[i(t)]$) in BA networks (bold) and ER network (thin) for outbreaks simulated on the same network (dashed curve), or on different networks (plain curve).The results are obtained for $\lambda = 0.01$ and on networks of size $N = 10^4$ nodes, and average degree $\langle k \rangle = 6$.

typical outbreak timescale τ given by [14, 15]

$$\tau = \frac{1}{\lambda(\kappa - 1)}. \tag{2}$$

A discussion of the validity of this equation is provided in Ref. [16]. In order to understand to which regime corresponds τ_v, we plot in Fig. 2 τ_v and τ for BA networks with different values of κ. We use networks with different sizes (from $N = 5.10^3$ to $N = 5.10^4$ nodes) and with different values of $\langle k \rangle$ ($6 < \langle k \rangle < 60$) in order to obtain a broad range of τ values.

We see in Fig. 2 that τ_v is increasing linearly with τ (with a pre-factor of order 4). This implies that τ_v is of the same order of the typical time τ_0 where the diversity of degree classes of infected nodes is the largest ($\tau_0 \approx 6 \tau$) [14, 15]. The result $\tau_v \approx \tau_0$ confirms the intuitive idea that the variability is maximal when the diversity of different classes of infected nodes is the largest, which happens at the beginning of the spread.

B. Effect of degree on $i(t)$ fluctuations

1. Seed degree

In this SI model, the parameter λ simply fixes the time unit. In contrast, we expect that other parameters such as the degree of the seed may have a more interesting effect on the outbreak and its variability. Fig. 3 displays the evolution of $CV[i(t)]$ for outbreaks starting from initial infected nodes with a given degree k_0 (from 3 up to 248). This figure shows that the variability peak decreases when k_0 is increased. In other words, when an outbreak begins from a highly connected node, the early stages of the spreading tend to be less variable. One might think that the number of paths available on a highly connected node leads to a higher overall variability, it is however not the case. As shown in the inset of Fig. 3, the prevalence increases with the seed degree, which may explain the variability for different k_0. Indeed, when the seed is a hub, the
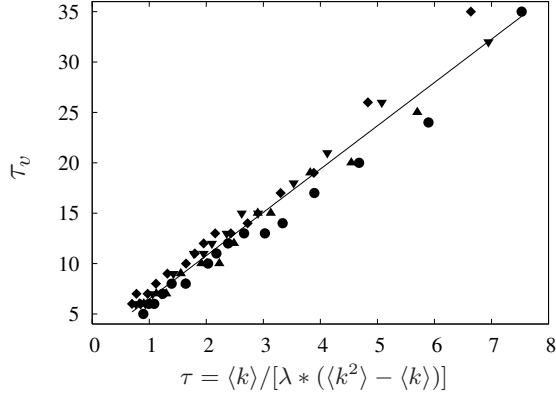
Figure 2: τ_v versus τ for several BA networks with $\langle k \rangle$ ranging from 6 to 60, and different sizes (●: $N = 5.10^3$, ▲: $N = 10^4$, ▼: $N = 2.10^4$, ◆: $N = 5.10^4$ nodes; $\lambda = 0.01$). The line is a linear fit with slope of order 4.
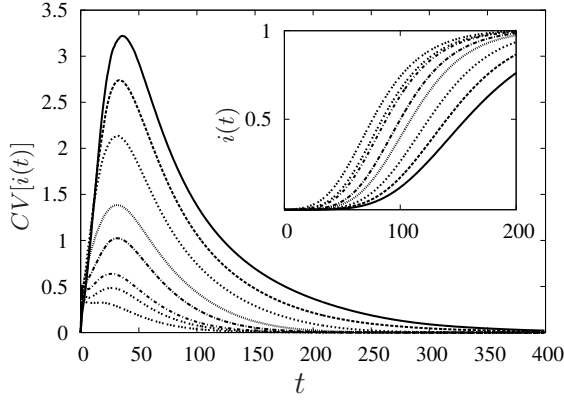

Figure 3: Temporal evolution of the coefficient of variation of the density of infected ($CV[i(t)]$) in BA networks for outbreaks seeded with infected nodes of different degrees k_0 (from top to bottom, $k_0 = 3, 6, 12, 24, 48, 95, 142, 248$). Inset: Initial evolution of the prevalence i(t). The order of the curves is reversed between both plots (Results are averaged over 5.10^3 epidemics on one network, with $\lambda = 0.01$, $N = 10^4$ nodes, $\langle k \rangle = 6$).

number of infected becomes rapidly very large and thus leads to smaller relative variations of the prevalence. This result leads us to investigate more thoroughly the degree of infected nodes and analyze the differences between BA and ER networks.

2. Degree of infected nodes

In this section, we study in detail the degree properties of the infected nodes during the outbreak of the disease.

For a SI model, the evolution of the density $i_k(t)$ of infected nodes of degree k is given at the mean-field level by

$$\frac{di_k(t)}{dt} = \lambda \, k \, [1 - i_k(t)] \, \Theta_k(t) \qquad (3)$$
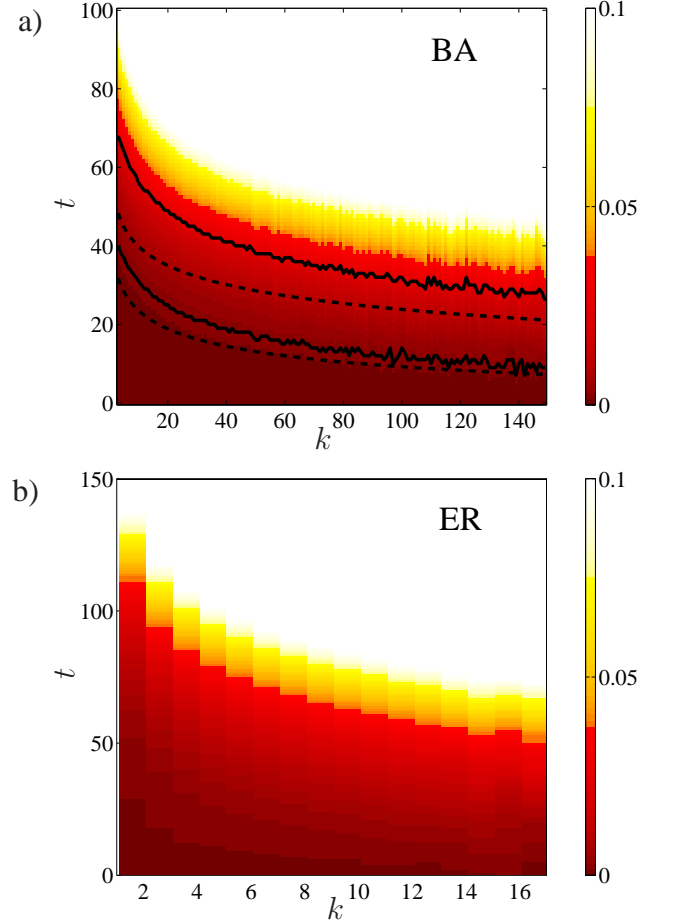

Figure 4: Temporal evolution of the density of infection by classes of degree. **(a)** Spreading on a BA network. The dashed lines are given by Eq. (5) with different values for i: 0.002, 0.02 (lower to upper) and the plain lines corresponds to their numerical results. **(b)** Spreading on an ER network. For both panels the color bar represents the density of infected and where white means 0.1 and above (The results are computed over 10^3 outbreaks on networks of size $N = 10^4$ nodes, $\langle k \rangle = 6$, and spreading rate $\lambda = 0.01$).

where $1 - i_k$ is the density of susceptible nodes of degree k and Θ_k is the probability that a link pointing to a node of degree k originates at an infected node [10]. This equation, studied for an uncorrelated scale-free network and uniform initial conditions $i_0 \equiv i_k(t = 0)$ leads to the following behavior at short times [14, 15]

$$i_k(t) \simeq i_0 \left[1 + \frac{k \, \langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} (e^{t/\tau} - 1) \right] \qquad (4)$$

with τ defined in Eq. (2).

From this equation, we can deduce the expression for the time $t_k(i)$ for i_k to reach the value i:

$$t_k(i) \simeq \tau \log \left[1 + \frac{\langle k^2 \rangle - \langle k \rangle}{k \langle k \rangle} \left(\frac{i}{i_0} - 1 \right) \right] \qquad (5)$$

For a fixed prevalence i, the time $t_k(i)$ varies very slowly with k and thus can vary significantly only on a network with a
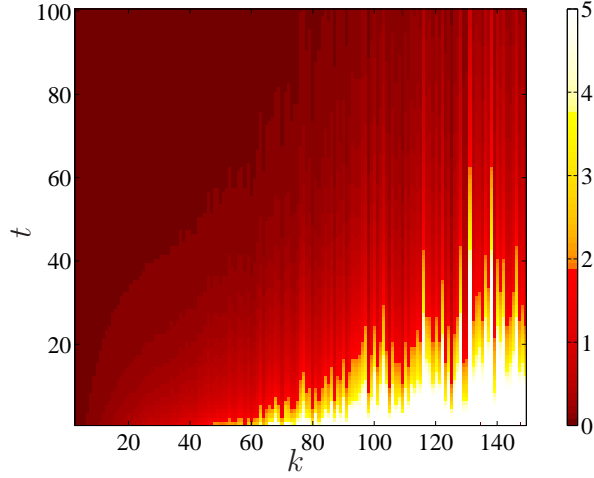
Figure 5: Temporal evolution of the coefficient of variation of the density of infected by classes of degree on a BA network. The range of $CV[i_k(t)]$ is limited to $[0, 5]$ for readability (the actual results can go up to 30). Color bar accounts for $CV[i_k(t)]$ (white means 10 and above). These results are obtained for $N = 10^4$ nodes, $\langle k \rangle = 6$, $\lambda = 0.01$, and averaged over 10^3 outbreaks on 50 different networks in order to have data for the whole range of degrees.

large range of degree variation. The results are shown on Fig. 4, which is composed of two contour maps of the temporal evolution of i_k in both BA and ER networks. In order to simplify the reading of this figure, the density of infection has been limited to 0.1 since we are only interested in the beginning of the outbreaks. We also plot, in Fig. 4(a), the curves corresponding to Eq. (5) for different values of i (0.002 and 0.02) and numerical result for the same values (plain curves). It can be seen that the predictions of Eq. (5) for small density and short times are in agreement with the average behavior obtained from our simulations (the agreement is better for larger degree since the hubs are infected at smaller times). For larger times, the approximation used in Eq. (4) is not valid anymore, which explains the observed discrepancy for larger values of the density such as $i = 0.02$. These results confirm earlier work [14, 15] on the "cascading effect" of the spreading, from hubs to poorly connected nodes. Figure 4(b) is the ER counterpart of Fig. 4(a). It demonstrates that the hierarchical spreading from well connected to poorly connected nodes also occurs on homogeneous networks. The cascading effect however is less visible on the average degree of infected nodes because of the limited range of degrees (see also Sec. III A).

Figure 5 gives a complete picture of the variability of $i_k(t)$ in an heterogeneous network and helps to understand the role of each degree in the variability peak observed in Fig. 1. It displays for a BA network a contour map representation of the temporal evolution of $CV[i(t)]$ according to the classes of degree. We observe that the largest values of $CV[i_k(t)]$ are reached at the beginning of outbreaks, then decrease during the infection process. The very high values of CV (white on the plot) which can be up to 30 are reached during a period lasting until 6 τ (in this plot $\tau \approx 7$). The end of this period corresponds to the moment when all degree classes are

infected. For superspreaders, Fig. 5 also shows that their infection time is fluctuating a lot even for long times, because of their small number in networks. This result will be confirmed in the next section and means that their infection time has important fluctuations. For some outbreaks, the time to reach a superspreader can be long because of its distance to the seed (see Sec. III B).

III. FLUCTUATIONS OF THE INFECTION TIME

The randomness of the epidemic process makes it very difficult to predict an accurate time interval for the infection of a given node. However, with the same methods used in the previous section, we can draw the general picture of the distributions of the infection time t_{inf}—defined as the time for which a given node becomes infected—as a function of the degree of the node and its distance to the seed (similar considerations were studied in [17]).

A. Effect of the degree

Fig. 4 shows how the prevalence $i(t)$ varies with the degree. Time of infection and prevalence being related, we first plot (Fig. 6) the distribution of the infection time t_{inf} versus the degree. For this figure we count all the nodes with a given degree k which have been infected at each instant t, and then we normalize the corresponding results by the number of individuals with degree k and by the number of simulations. Each degree is represented by a column where frequencies are associated with a representative color (right color bar), the sum of all frequencies in a column being equal to one. Given that a single BA network does not contain the whole range of degrees, the plot shown on Fig. 6(a) is based on data from 50 networks. These results are a consequence of the cascading effect on lower degree nodes on both topologies: the larger the degree and the smaller the average infection time. In addition, we observe that there is a relatively large range of fluctuation of the infection time even for large degrees. Indeed, in the inset of Fig. 6(a) we observe that for highly connected nodes (e.g. from 80 to 150), the typical t_{inf} varies between 6 τ and 13 τ (on the plot, $t = 40$ and 90) which is late for well-connected nodes. In fact, only a small fraction of the superspreaders is infected during the early epidemic stages (until 6 τ) and triggers the outbreak. Approximately the same scenario seems to hold for ER networks (Fig. 6(b)), even if the concept of superspreaders is not the most appropriate for a network with a small range of degree variation.

In order to understand thoroughly the properties of the infection time, we also show in Fig. 7 scatter-plots of its relative dispersion $CV(t_{inf})$ versus the degree for both ER and BA topologies. This figure displays more insights concerning the behavior of t_{inf} depicted in Fig. 6. First, for the BA network, nodes with a given degree k can have a wide range of $CV(t_{inf})$ which increases with k. This demonstrates that even if the superspreaders are infected at relatively short times, large relative fluctuations cannot be excluded. In con-
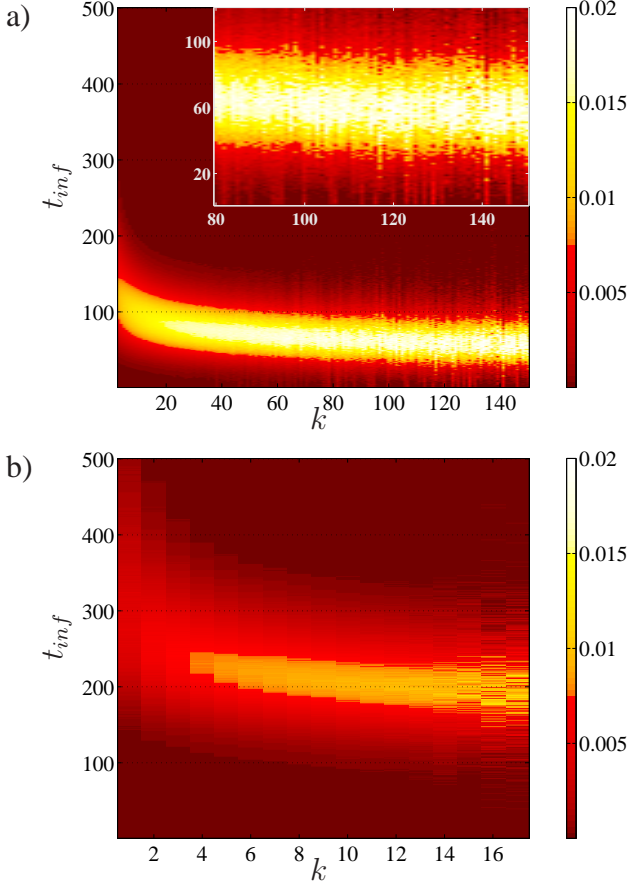
Figure 6: Frequency of the moment of infection of a node as a function of its degree. **(a)** spreading on a BA network. Inset: frequency of t_{inf} shown for the beginning of outbreaks on high degree nodes. **(b)** spreading on an ER network. For both panels the color bar represents the frequency and where white means 0.02 and above. ($N = 10^4$ nodes, $\langle k \rangle = 6$, $\lambda = 0.01$, $\tau_{BA} \approx 7$, $\tau_{ER} \approx 16.5$).

trast, all nodes for the ER network have smaller and similar values of $CV(t_{inf})$ which is consistent with the fact that the hierarchical spreading is less pronounced on ER due to its limited range of degree.

B. Effect of distance

Another important parameter which affects the infection time of a node is its distance to the seed as measured by the number of hops of the shortest path [17]. In the networks considered here there is no spatial component and the distance between two nodes is given by the smallest number of hops ℓ to go from one node to another. On Fig. 8, we show the relationship between the average time of infection $\langle t_{inf} \rangle$ and ℓ for ER and BA networks. We see on this plot that the infection time $\langle t_{inf} \rangle$ is always larger for ER than for BA networks. It means that nodes with the same value of ℓ, i.e. at the same distance from the first infected node, have a lower $\langle t_{inf} \rangle$ if they belong to a BA networks. The reason for this behavior
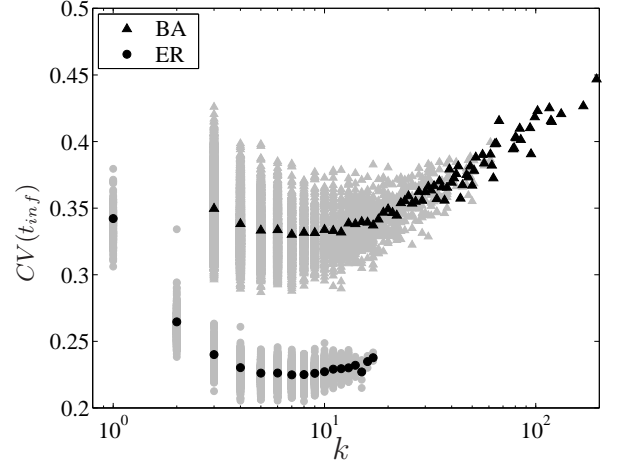

Figure 7: Coefficient of variation of infection time as a function of nodes degree k. Gray symbols stand for $CV(t_{inf})$ computed by nodes, and black symbols for $CV(t_{inf})$ computed for nodes with the same degree (vertically aligned). ▲ symbols represent the spread on BA networks, and ● stand for ER networks (Results are computed over 10^3 outbreaks on a single network, $N = 10^4$ nodes, $\langle k \rangle = 6$ links, with a seed of degree $k_0 = 6$, $\lambda = 0.01$).
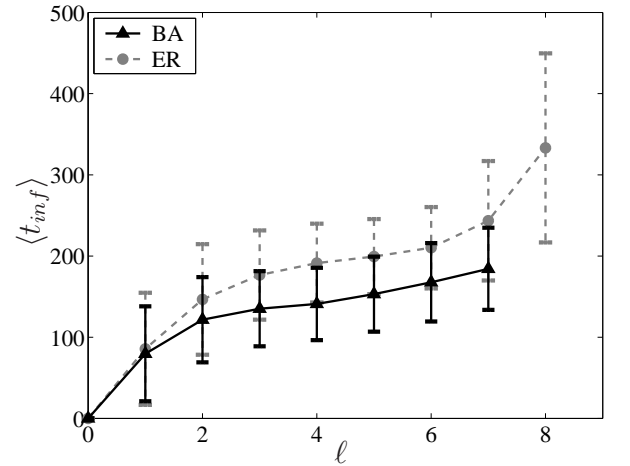

Figure 8: Infection time as a function of the distance ℓ from the seed ($N = 10^4$, $\langle k \rangle = 6$, averaged over 10^3 outbreaks which start at exactly the same seed of degree $k_0 = 6$).

ℓ	BA networks	ER networks
1	1	1
2	1.00668	1.00144
3	1.09870	1.01126
4	1.48679	1.06732
5	2.44046	1.40124
6	3.25166	2.77646
7	3.15678	2.64469
8	–	1.11256
9	–	1

Table I: Average number of shortest paths between a randomly selected node and a node at distance ℓ (results are computed over 10^3 random selections of an initial node of degree $k_0 = 6$). $N = 10^4$ nodes, $\langle k \rangle = 6$.

lies in the difference of the numbers of shortest paths in these networks. Indeed, if we enumerate these paths, we observe that their numbers relatively differ between both BA and ER topologies. We have computed the size and the number of shortest paths between a randomly selected node, i.e. a potential seed of infection and the rest of the network and we present in Table I the average number of shortest path at distance ℓ. Results are computed over 10^3 random selection of the potential seed in order to get an accurate picture of the network. The table exhibits a difference in the number of path for $\ell > 2$ (8% difference for $\ell = 3$, 40% for $\ell = 4$, 74% for $\ell = 5$) which confirms the fact that on BA networks, nodes have more paths to go from one to another in a small number of hops. Table I describes the statistics of shortest paths but
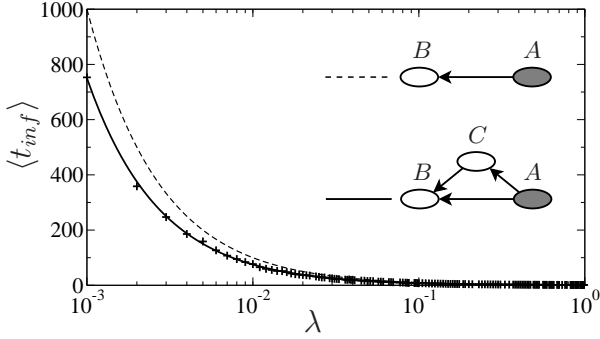

Figure 9: Average infection time of node B as a function of λ for two different configurations. In the first case infection occurs in one step and in the second case another path is added. The dotted curve represents the average time of infection for the first case, $\langle t_{inf} \rangle = 1/\lambda$ and the plain curve represents $\langle t_{inf} \rangle$ for the second case and is given by Eq. (7). The result of a numerical simulation are shown by + symbols).

longer paths also contributes to the spreading of the disease. Their role can be highlighted by studying the following simple cases. In the first case an infected node A is in contact with a susceptible node B. In the second case, there is an additional path from A to B going through a susceptible node C (see Fig. 9). In the first "direct" case, the average time of infection $\langle t^d_{inf} \rangle(B)$ of B is given by

$$\langle t^d_{inf}(B) \rangle = \frac{1}{\lambda}. \tag{6}$$

The addition of a longer path in the second case (Fig. 9) changes the behavior of $\langle t_{inf}(B) \rangle$ and Eq. (6) no longer holds for this case. In fact, the time of infection of the susceptible node B is given by

$$t_{inf}(B) = \min[t^d_{inf}(B), t^i_{inf}(B)], \tag{7}$$

where $t^d_{inf}(B)$ is the time of a direct infection $A \rightarrow B$ and t^i_{inf} of an indirect 2-steps infection process: $A \rightarrow C \rightarrow B$. The statistics of t_{inf} can be easily computed and its first moment reads

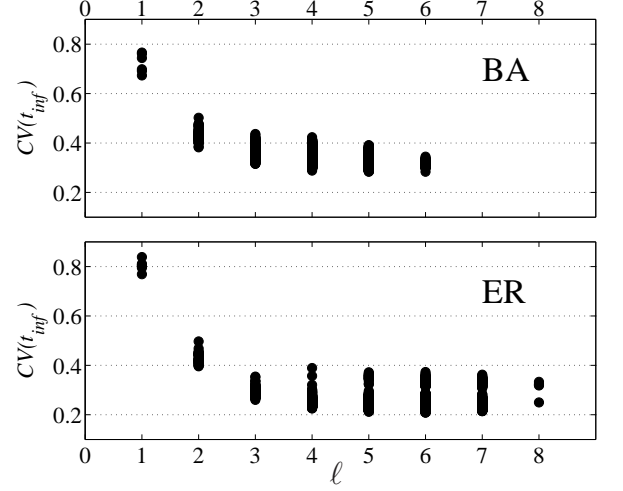$$\langle t_{inf}(B) \rangle = \frac{1}{\lambda} \frac{3 - 2\lambda}{(2 - \lambda)^2} \tag{8}$$

Figure 10: Coefficient of variation of the infection time as a function of the distance ℓ from the seed. Top panel: spreading on a BA network; bottom panel: ER network. Both panels show $CV(t_{inf})$, for every nodes of a single network, $N = 10^4$ nodes, $\langle k \rangle = 6$, and computed over 10^3 outbreaks, $\lambda = 0.01$, originating from exactly the same seed of degree $k_0 = 6$.

Eq. (8) predicts values always smaller than $1/\lambda$ (see Fig. 9). This result could appear as paradoxical since adding a *longer* path actually *reduces* the average infection time. In fact, the probability that the disease is not transmitted on both paths is very small and the existence of another path cuts off large direct infection time and thus reduces the average infection time of B. Since BA networks have a clustering coefficient larger than ER networks [1] this result explains the small difference of infection times for $\ell = 1$ seen in Fig. 8.

Concerning the relationships between the relative dispersion of infection time $CV(t_{inf})$ and ℓ, their behavior on both topologies are reported on Fig. 10. This figure shows that the nodes in both networks exhibit higher values of $CV(t_{inf})$ when they are closer to the seed, i.e. for $\ell < 3$. For larger distances, $CV(t_{inf})$ is practically constant in both cases.

IV. CONCLUSIONS

We have analyzed in detail the variability of a simple epidemic process on SFN. First, we have shown that different realizations of BA networks do not display significant statistical differences in outbreak variability. Consequently, it is statistically reliable to consider a single realization of the network, provided it is large enough. We have also shown that the prevalence fluctuations are maximal during the time regime for which the diversity of the degrees of the infected node is the largest. In order to analyze in detail this variability, we examined the temporal degree pattern of infected nodes. In particular, we demonstrated the high variability of superspreaders' prevalence. We found that for the hubs the infection time is usually small but with fluctuations which can be large. Even if the hubs are good candidates for being chosen as surveil-

lance stations—given their short average infection time, they present non-negligible fluctuations which limit their reliability. In this respect, the ideal detection stations should be nodes with the best trade-off between a short average infection time and a high reliability as given by small infection time fluctuations.

The topological distance to the seed is also an important parameter in epidemic spreading pattern. Nodes at a short distance from the seed are infected at small time—in the high variability regime—and thus have a large infection time variability. Maybe more surprising is the importance of the number of paths—not only the shortest one—going from the seed to another node. The larger this number and the smaller the average infection time. This is an important conclusion for containment strategies since the reduction of epidemic channels will increase the delay of the infection arrival and will thus allow for a better preparation against the disease (for example vaccination).

These results could be helpful in designing early detection and containment strategies in more involved models which go beyond topology and which include additional features such as passenger traffic in airlines or city populations [20, 29, 30, 31].

Acknowledgments

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2000).

[2] S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of networks: From biological nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).

[3] R. Pastor-Satorras and A. Vespignani, *Evolution and structure of the Internet: A statistical physics approach* (Cambridge University Press, Cambridge, 2003).

[4] M.E.J. Newman, SIAM Review **45**, 167 (2003).

[5] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[6] L.A.N. Amaral, A. Scala, M. Barthélemy, and H.E. Stanley, Proc. Natl. Acad. Sci. USA **97**, 11149 (2000).

[7] F. Liljeros, C. R. Edling, L.A.N. Amaral, H.E. Stanley, and Y. Aberg, Nature **411**, 907 (2001).

[8] A. Schneeberger, C.H. Mercer, S.A. Gregson, N.M. Ferguson, C.A. Nyamukapa, R.M. Anderson, A.M. Johnson, and G.P. Garnett, Sex. Transm. Dis. **31**, 380 (2004)

[9] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)

[10] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001)

[11] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 66117 (2001).

[12] R.M. May and A.L. Lloyd, Phys. Rev. E **64**, 66112 (2001).

[13] A.L. Lloyd and R.M. May, Science **292**, 1316 (2001).

[14] M. Barthélemy, A. Barrat, R. Pastor-Satorras, A. Vespignani, Phys. Rev. Lett. **92**, 178701 (2004).

[15] M. Barthélemy, A. Barrat, R. Pastor-Satorras, A. Vespignani, J. Theor. Biol., **235**, 275 (2005).

[16] A. Vazquez, Phys. Rev. Lett. **96**, 038702 (2006).

[17] M. Loecher and J. Kadtke, submitted (2005).

[18] H.W. Hethcote and J.A. Yorke, *Gonorrhea Transmission Dynamics and Control*, Lect. Notes Biomath. **56**, 1-105 (Springer, Berlin, 1984).

[19] R.M. Anderson and R.M. May, *Infectious diseases in humans. Dynamics and control* (Oxford University Press, Oxford, 1992).

[20] V. Colizza, A. Barrat, M. Barthélemy, A. Vespignani, Proc. Natl. Acad. Sci. U.S.A. **103**, 2015 (2006)

[21] O. Diekmann and J.A.P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation* (Wiley & Sons, 2000).

[22] M.J. Keeling and B.T. Grenfell, J. Theor. Biol. **203**, 51 (2000).

[23] M. Lipsitch *et al*, Science **300**, 1966 (2003).

[24] S. Riley *et al*, Science **300**, 1961 (2003).

[25] A.P. Galvani and R.M. May, Nature **438**, 293 (2005).

[26] J.O. Lloyd-Smith, S.J. Schreiber, P.E. Kopp, and W.M. Getz, Nature **438**, 355 (2005).

[27] P. Erdös and A. Renyi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).

[28] M. Boguñá, R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **90**, 28701 (2003).

[29] A. Flahault and A.-J. Valleron, Math. Pop. Studies. **3**, 1 (1991).

[30] L. Hufnagel, D. Brockmann, T. Geisel, Proc. Natl. Acad. Sci. U.S.A. **101**, 15124 (2004).

[31] S. Eubank, H. Guclu, V.S. Anil Kumar, M.V. Marathe, A. Srinivasan, Z. Toroczkai, N. Wang, Nature **429**, 180 (2004).